

IN THE UNITED STATES DISTRICT COURT  
FOR THE DISTRICT OF DELAWARE

BRITISH TELECOMMUNICATIONS PLC, )

Plaintiff, )

v. )

GOOGLE INC., )

Defendant. )

C. A. No.: 11-1249-LPS

**AGREEMENT AND [PROPOSED] ORDER REGARDING  
THE FORMAT OF DOCUMENT PRODUCTION**

This Agreement and Order Regarding the Format of Document Production (“Agreement”) supplements and modifies all other discovery rules, agreements, and orders and shall govern the parties in the above-captioned case (the “Litigation”). This Agreement may be modified in the Court’s discretion or by agreement of the parties. If the parties cannot resolve their disagreements regarding such modifications, the parties shall submit their competing proposals and a summary of their dispute pursuant to the Court’s applicable practices and procedures.

**I. GENERAL PROVISIONS**

**A. PRODUCTION.** The parties shall make reasonable efforts to prepare responsive and nonprivileged data for production in accordance with the agreed-upon specifications set forth below. These specifications apply to hard copy documents and to electronically stored information (“ESI”) which are to be produced in the first instance in this Litigation.

**B. SECURITY.** Both parties shall make reasonable efforts to ensure that any productions made are free from viruses and provided on encrypted media for submission.

**C. CONFIDENTIALITY DESIGNATION.** Responsive documents in TIFF format shall be stamped with the appropriate confidentiality designations in accordance with the

Agreed Protective Order in this matter. Each responsive document produced in native format shall have its confidentiality designation identified in the filename of the native file.

**D. NON-STANDARD FILES.** The parties shall discuss the format of production of non-standard electronic files, large oversized documents, etc. before production to determine the optimal production format.

**E. PRODUCTION MEDIA.** Documents shall be produced on external hard drives, readily accessible computer(s) or electronic media (collectively, "Production Media"). Each piece of Production Media shall identify a production number corresponding to the production volume (e.g., "VOL001", "VOL002"), as well as the volume of the material in that production (e.g. "-001", "-002"). Each piece of Production Media shall also identify: (1) the producing party's name; (2) the production date; (3) the Bates Number range of the materials contained on the Production Media; and (4) the set(s) of requests for production for which the documents are being produced.

## **II. DATA PROCESSING**

**A. KEYWORD SEARCHING.** To the extent that keywords are used in limiting the universe of potentially responsive documents to be reviewed, the parties shall meet and confer to try to develop a mutually agreeable list of search terms and protocols prior to the production of documents. Any search methodology employed must open compound and nested files and de-compress archived files. The search utilities employed must support the use of stemmed searches (e.g. using "~" to include variations on a keyword) and Boolean searches.

**B. CULLING/FILTERING.** Each party shall use its best efforts to filter out common system files and application executable files by using a commercially reasonable

hash identification process. Hash values that may be filtered out during this process are located in the National Software Reference Library (“NSRL”) hash set list. Additional culling of system file types based on file extension may include, but are not limited to: WINNT, LOGS, DRVS, MP3, C++ Program File (c) , C++ Builder 6 (cpp), Channel Definition Format (cdf), Creatures Object Sources (cos), Dictionary file (dic), Executable (exe), Hypertext Cascading Style Sheet (css), JavaScript Source Code (js), Label Pro Data File (IPD), Office Data File (NICK), Office Profile Settings (ops), Outlook Rules Wizard File (rwz), Scrap Object, System File (dll), Temporary File (tmp), Windows Error Dump (dmp), Windows Media Player Skin Package (wmz), Windows NT/2000 Event View Log file (evt), Python Script files (.py, .pyc, .pud, .pyw), Program Installers. Notwithstanding the foregoing, the parties reserve the right to request that certain system files be produced.

**C. ACCESS TO PROTECTED OR ENCRYPTED FILES.** The parties agree to use commercially accessible software to open password-protected or encrypted files to the extent that such files can be opened. If certain password-protected or encrypted files cannot be opened, the parties will meet and confer to address how to produce such files.

**D. DEDUPLICATION.** A party is only required to produce a single copy of a responsive document and a party may de-duplicate responsive ESI (based on MD5 or SHA-1 hash values at the document level) across custodians. For emails with attachments, the hash value is generated based on the parent/child document grouping. A party may also de-duplicate “near-duplicate” email threads as follows: In an email thread, only the final-in-time document need be produced, assuming that all previous emails in the thread are contained within the final message. Where a prior email contains an attachment, that email and attachment shall not be removed as a “near-duplicate.” To the extent that

de-duplication through MD5 or SHA-1 hash values is not possible, the parties shall meet and confer to discuss any other proposed method of de-duplication. Notwithstanding the foregoing, the parties reserve the right to request that all copies of a particular document be produced.

### **III. PRODUCTION OF HARD COPY DOCUMENTS**

**A. TIFFs.** Documents that exist in hard copy format only shall be scanned and produced as single page Group IV TIFFs, with at least 300 dots per inch (dpi). Each TIFF image shall be named according to the corresponding Bates number associated with the document. Each image shall be branded according to the Bates number and agreed upon confidentiality designation. TIFFs shall show all text and images that would be visible to a user of the hard copy documents. If a color copy is reasonably necessary to comprehend a document and a receiving party requests a color copy, such documents will be produced as single-page JPEGs.

**B. OCR TEXT FILES.** A commercially acceptable technology for optical character recognition (“OCR”) shall be used for all scanned, hard copy documents. OCR text shall be provided as a single text file for each document, and the filename itself should match its respective TIFF filename. The text files shall not contain the redacted portions of the documents.

**C. DATABASE LOAD FILES/CROSS-REFERENCE FILES.** Documents shall be provided with (a) an ASCII delimited data file (.txt, .dat, or .csv), and (b) an image load file that can be loaded into commercially acceptable production software (e.g., Concordance, Summation). Each TIFF in a production must be referenced in the corresponding image load file. The total number of documents referenced in a

production's data load file should match the total number of designated document breaks in the image load file(s) in the production.

**D. CODING FIELDS.** The following information shall be produced in the delimited data file accompanying hard copy documents: (a) BEGBATES, (b) ENDBATES, (c) CUSTODIAN, (d) CONFIDENTIALITY, and (e) REDACTED.

**E. BATES NUMBERING.** All images must be assigned a unique Bates number that is sequential within a given document and across the production sets.

**F. UNITIZING OF DOCUMENTS.** In scanning paper documents, distinct documents should not be merged into a single record, and single documents should not be split into multiple records (*i.e.*, paper documents should be logically unitized). The parties shall use reasonable efforts to unitize documents correctly.

**G. IDENTIFICATION OF PAPER DOCUMENTS.** The parties shall use best efforts to ensure that paper records for a particular custodian that are included in a single production are produced in consecutive Bates stamp order. The parties shall identify which documents in a production are scanned paper documents either in the cover letter accompanying the production or in a coding field titled "Paper (Y/N)".

#### **IV. PRODUCTION OF ELECTRONICALLY STORED INFORMATION ("ESI")**

**A. METADATA FIELDS AND PROCESSING.** Each of the metadata and coding fields set forth in Appendix 1 that can be extracted shall be produced for that document. The parties are not obligated to populate manually any of the fields in Appendix 1 if such fields cannot be reasonably extracted from a document, with the exception of the following: (a) BEGBATES, (b) ENDBATES, (c) BEGATTACH, (d) ENDATTACH; (e) CUSTODIAN, (f) CONFIDENTIALITY, and (g) REDACTED, which should be

populated by the party or the party's vendor. The parties shall make reasonable efforts to ensure that metadata fields automatically extracted from the documents are correct; however, the parties acknowledge that such metadata may not always be accurate and might instead contain irrelevant or incorrect information generated during the collection process. Parties may request other native files be produced as described in Section IV.I. below.

**B. TIFFs.** Unless excepted below, single page Group IV TIFFs shall be provided, at least 300 dpi. Each TIFF image file shall be one page and named according to the unique Bates number, followed by the extension ".TIF". Original document orientation shall be maintained (i.e., portrait to portrait and landscape to landscape). If a color copy is reasonably necessary to comprehend a document and a receiving party requests a color copy, such documents will be produced as single-page JPEGs.

**C. TEXT FILES.** For each document, a text file shall be provided along with the image files and metadata. The text of native files shall be extracted directly from the native file. However, if a document has been redacted or does not contain extractable text, OCR of the redacted document will suffice in lieu of extracted text.

**D. DATABASE LOAD FILES/CROSS-REFERENCE FILES.** Each TIFF in a production shall be produced with an appropriate database load file or cross-reference file, including (a) an ASCII delimited data file (.txt, .dat, or .csv), and (b) Concordance and Opticon load files that can be loaded into commercially acceptable production software. Each TIFF in a production must be referenced in the corresponding image load file. The total number of documents referenced in a production's data load file shall match the total number of designated document breaks in the image load file(s) in the production. The

total number of documents in a production shall match the total number of records in the database load file. Each TIFF in a production must be referenced in the corresponding image load file. The total number of pages referenced in a production's image load file shall match the total number of TIFF files in the production.

**E. BATES NUMBERING.** All images must be assigned a unique Bates number that is sequential within a given document and across the production sets.

**F. PRESENTATIONS.** The parties shall take reasonable efforts to process presentations (e.g., MS PowerPoint, Google Presently/Punch) with hidden slides and speaker's notes unhidden, and to show both the slide and the speaker's notes on the TIFF image.

**G. SPREADSHEETS.** TIFF images of spreadsheets need not be produced unless redacted, in which instance, spreadsheets shall be produced in TIFF with OCR. Native copies of spreadsheets shall be produced with a link in the NativeLink field, along with extracted text and applicable metadata fields set forth in Appendix 1. A TIFF placeholder indicating the document was provided in native format shall accompany the database record. If a spreadsheet has been redacted, TIFF images and OCR of the redacted document shall suffice in lieu of a native file and extracted text. The parties shall make reasonable efforts to ensure that any spreadsheets that are produced only as TIFF images are formatted so as to be readable.

**H. PROPRIETARY FILES.** To the extent a response to discovery requires production of ESI accessible only through proprietary software, the parties shall continue to preserve each version of such information. The parties shall meet and confer to finalize the appropriate production format.



**I. REQUEST(S) FOR ADDITIONAL NATIVE FILES.** If good cause exists to request production of specified files, other than those specifically set forth above, in native format, the party may request such production and provide an explanation of the need for native file review, which request shall not unreasonably be denied. Any native files that are produced shall be produced with a link in the NativeLink field, along with all extracted text and applicable metadata fields.

**J. REDACTION OF INFORMATION.** If documents are produced containing redacted information, the producing party shall supply a list of the documents for any such claim(s) of privilege, indicating the grounds for the redaction and the nature of the redacted material (e.g., privilege, privacy). During the course of this Litigation, an electronic copy of the originally unredacted data shall be securely preserved in such a manner so as to preserve without modification, alteration or addition the content of such data including any metadata therewith. If scanning is necessary to create a redacted version of a document, the producing party shall produce a corresponding load file identifying the custodian from whom the document was collected and such documents shall subject the TIFF files to OCR process if the TIFF files do not already have extractable text, such that the processed files become searchable.

**K. PRIVACY.** The parties acknowledge that some of the data may be subject to foreign data protection laws, and to the extent that a producing party agrees to produce any such data, the parties agree to confer on how to produce such data, including the extent of appropriate redactions.



## **V. PROCESSING OF THIRD-PARTY DOCUMENTS**

- A.** A party that issues a non-party subpoena (“Issuing Party”) shall include a copy of this Agreement with the subpoena and state that the parties to the Litigation have requested that third-parties produce documents in accordance with the specifications set forth herein.
- B.** The Issuing Party shall produce any documents obtained pursuant to a non-party subpoena to the opposing party.
- C.** If the non-party production is not Bates-stamped, the Issuing Party shall endorse the non-party production with unique prefixes and Bates numbers prior to producing them to the opposing party.
- D.** Nothing in this Agreement is intended to or should be interpreted as narrowing, expanding, or otherwise affecting the rights of the parties or third-parties to object to a subpoena.

## **VI. SEARCHING**

- A. CUSTODIAL PRODUCTION.** The parties shall meet and confer to develop a list of custodians and search terms with which the parties will conduct electronic searches using such agreed search term lists of custodial documents. The parties may jointly agree to modify such limits without the Court’s leave.

Search terms shall be employed for the searching of custodial data. The total search terms shall be narrowly tailored to particular issues and mutually agreed upon after each party performs search term analytics to optimize which set of search terms to employ (or as otherwise modified by agreement by the parties). To the extent that the parties are unable to mutually agree upon search terms, the parties shall employ the search methodology pursuant to Section 5.b. of the Delaware Default Standard for Discovery. A conjunctive

combination of multiple words or phrases (e.g., “computer” and “system”) narrows the search and shall count as a single search term. A disjunctive combination of multiple words or phrases (e.g., “computer” or “system”) broadens the search, and thus shall count as a separate search term unless they are variants of the same word.

**B. LOCATIONS THAT SHALL BE SEARCHED FOR RESPONSIVE DOCUMENTS.** The parties shall search any electronic files or folders, or other parts of media, including any internal and external hard drives and other ESI venues (including, but not limited to, recordable optical media, media cards, thumb drives, non-volatile memory, floppy disks, work desktop and laptop computers, email servers, intranet servers, network shares, public data shares and/or database servers) for each identified custodian that the custodian reasonably anticipates may contain non-duplicative responsive documents.

The parties agree to search central repositories, including central databases, or relevant portions thereof to the extent that the party reasonably anticipates they contain non-duplicative responsive documents. The parties agree to meet and confer to limit the scope of production from central repositories if the search of central repositories (or relevant portions thereof) that the producing party anticipates contain responsive documents is unduly burdensome or is likely to be unreasonably inefficient in identifying relevant documents. Specifically, the parties recognize that certain repositories, by their nature, may not effectively or reasonably be searched using electronic search strings, and the parties agree to notify each other of any such repositories that contain responsive documents. The parties will then meet and confer to discuss the collection of responsive documents from such repositories, including potentially using other effective collection methodologies.

**C. LOCATIONS THAT SHALL NOT BE SEARCHED FOR RESPONSIVE DOCUMENTS.** The categories of ESI identified in Appendix 2 shall not be searched under any circumstances, and as such need not be preserved. In addition, the parties agree that only sent and received custodial emails shall be searched. Notwithstanding the foregoing, the parties agree that responsive documents that a custodian indicates are stored on an archival storage medium that the custodian can readily identify and locate, that cannot be located in any other repository of information, and that can reasonably be searched, shall be searched. In addition, nothing in this paragraph shall limit a receiving party's right to request from a producing party more information about the nature of and burden associated with obtaining documents from a particular location. The parties further recognize their obligations to preserve any potentially relevant sources of data, whether live or in archival form, for purposes of this Litigation.

**D. SOURCE CODE.** To the extent relevant to the Litigation, source code shall be made available for inspection pursuant to the terms of the Agreed Protective Order. The parties agree that the search terms shall not be applied to source code.

## **VII. MISCELLANEOUS PROVISIONS**

**A. OBJECTIONS PRESERVED.** Nothing in this Agreement shall be interpreted to require disclosure of irrelevant information or relevant information protected by the attorney-client privilege, work-product doctrine, or any other applicable privilege or immunity. The parties do not waive any objections as to the production, discoverability, admissibility, or confidentiality of documents and ESI.

**B. NO EFFECT ON COST SHIFTING.** Nothing in this Agreement shall affect, in any way, a producing party's right to seek reimbursement for costs associated with collection, review, and/or production of documents or ESI.

**C. NO WAIVER.** Pursuant to Federal Rule of Evidence 502(d), the inadvertent production of a privileged or work product protected ESI is not a waiver in the pending case or in any other federal or state proceeding. Moreover, the mere production of ESI in a litigation as part of a mass production shall not itself constitute a waiver for any purpose.

MORRIS, NICHOLS, ARSHT & TUNNELL  
LLP

/s/ Paul Saindon

Jack B. Blumenfeld (No. 1014)  
Paul Saindon (No. 5110)  
1201 North Market Street  
P.O. Box 1347  
Wilmington, DE 19899  
(302) 658-9200  
jblumenfeld@mnat.com  
psaindon@mnat.com

*Attorneys for Google Inc.*

OF COUNSEL:

Dimitrios T. Drivas  
Kevin X. McGann  
John P. Scheibeler  
Robert E. Counihan  
WHITE & CASE LLP  
1155 Avenue of the Americas  
New York, NY 10036  
(212) 819-8200

Dated: August 22, 2012

YOUNG CONAWAY STARGATT &  
TAYLOR, LLP

/s/ Monté T. Squire

Melanie K. Sharp (No. 2501)  
Monté T. Squire (No. 4764)  
Robert M. Vrana (No. 5666)  
Rodney Square  
1000 North King Street  
Wilmington, Delaware 19801  
P.O. Box 391  
Wilmington, Delaware 19899-0391  
(302) 571-6681  
msharp@ycst.com

PROSKAUER ROSE LLP

James H. Shalek (admitted *pro hac vice*)  
Robert S. Mayer (admitted *pro hac vice*)  
Joon R. Yoon (admitted *pro hac vice*)  
Fabio E. Tarud (admitted *pro hac vice*)  
Eleven Times Square  
New York, New York 10036  
(212) 969-3000

PROSKAUER ROSE LLP

Justin J. Daniels (admitted *pro hac vice*)  
Sean D. Whaley (admitted *pro hac vice*)  
One International Place  
Boston, MA 02110-2600  
(617) 526-9600

*Attorneys for British Telecommunications plc*

SO ORDERED this \_\_\_\_ day of \_\_\_\_\_, 2012.

\_\_\_\_\_  
UNITED STATES DISTRICT JUDGE

**Appendix 1: ESI Metadata and Coding Fields**

**A. Image Load File** shall contain the following comma-delimited fields:  
BEGBATES, VOLUME, IMAGE FILE PATH, DOCUMENT BREAK, FOLDER  
BREAK, BOX BREAK, PAGE COUNT

**B. Metadata Load File** shall be delimited according to the following characters:  
o Delimiter = D(ASCII:0020)  
o Text-Qualifier = p (ASCII:00254)

**C.** The following fields will appear in the metadata load file in the order displayed below:

Field Name	Field Description
BEGBATES	Beginning Bates number as stamped on the production image
ENDBATES	Ending Bates number as stamped on the production image
BEGATTACH	First production Bates number of the first document in a family
ENDATTACH	Last production Bates number of the last document in a family
CUSTODIAN	Individual from whom the documents originated
NATIVELINK	Native File Link (Excel files only)
SUBJECT	Subject line of email
DATESENT	Date email was sent (format: MM/DD/YYYY)
TIMESENT	Time email was sent
TO	All recipients that were included on the "To" line of the email
FROM	The name and email address of the sender of the email
CC	All recipients that were included on the "CC" line of the email
BCC	All recipients that were included on the "BCC" line of the email
AUTHOR	Any value populated in the Author field of the document properties
FILENAME(Edoc only)	Filename of an electronic document
DATEMOD (Edoc only)	Date an electronic document was last modified (format: MM/DD/YYYY)
DATECREATED (Edoc only)	Date the document was created (format: MM/DD/YYYY)
FILEEXTENSION	The file extension of a document.
FILESIZE	The file size of a document (including embedded attachments)
MD5HASH	The MD5 Hash value, or "de-duplication key," assigned to a document.

## **Appendix 2**

1. Deleted, slack, fragmented, or other data only accessible by forensics.
2. Random access memory (RAM), temporary files, or other ephemeral data that are difficult to preserve without disabling the operating system.
3. On-line access data such as temporary internet files, history, cache, cookies, and the like.
4. Data in metadata fields that are frequently updated automatically, such as last-opened dates.
5. Back-up data that are substantially duplicative of data that are more accessible elsewhere, including automated disaster recovery backup systems and disaster recovery backup tapes.
6. Voice messages, including voicemail systems.
7. Instant messages that are not ordinarily printed or maintained in a server dedicated to instant messaging. For the avoidance of doubt, this category of ESI does not include instant messages that are maintained and saved in email servers to the extent maintained in the ordinary course of business (*e.g.*, instant messages exchanged via Google Talk, Google Chat, Microsoft Lync and Microsoft Office Communicator that are stored or maintained on email servers in the ordinary course of business).
8. Electronic mail or pin-to-pin messages sent to or from mobile devices (*e.g.*, iPhones, Blackberries, and personal data assistants), provided that a copy of such mail is routinely saved elsewhere.
9. Other electronic data stored on a mobile device, such as calendar or contact data or notes, provided that a copy of such information is routinely saved elsewhere.
10. Logs of calls made from mobile devices.
11. Server, system or network logs.
12. Electronic data temporarily stored by laboratory equipment or attached electronic equipment, provided that such data is not ordinarily preserved as part of a laboratory report.
13. Data remaining from systems no longer in use that is unintelligible on the systems in use.
14. Data remaining from systems no longer in use that are substantially duplicative of data that are more accessible elsewhere.